The objectives of the BF Geocoding Quality Paper were to:

- promote the sharing of information about the current potential uses of geocoding business frames.
- identify and explain the existing set of uses being made of geocoding.
- measure the existing quality of geocoding on the SNZ Business Frame.
- describe issues regarding the quality of addresses for geocoding of physical locations and techniques for improving address quality.
- identify components of SNZ current practices that may have created deficiencies and discuss possible remedial action.

What is Geocoding?

Geocoding is the process of creating geographic co-ordinates for geographically referenced tabular data. A geocoding process will allow the derivation of precise co-ordinates on the surface of the earth for things like physical addresses.
Generally there are two ways in which geocoding can be employed for the collection and presentation of statistical information. The first of these methods is a polygon-based approach where statistics for small and large scale areas are linked to polygons overlaid onto a topographic background map. The second method involves the use of a point-based approach where the statistical information is linked to x/y co-ordinates. Statistics New Zealand (SNZ) uses a polygon based approach.

New Zealand is divided geographically into 16 regional areas. These are then subdivided into the following :

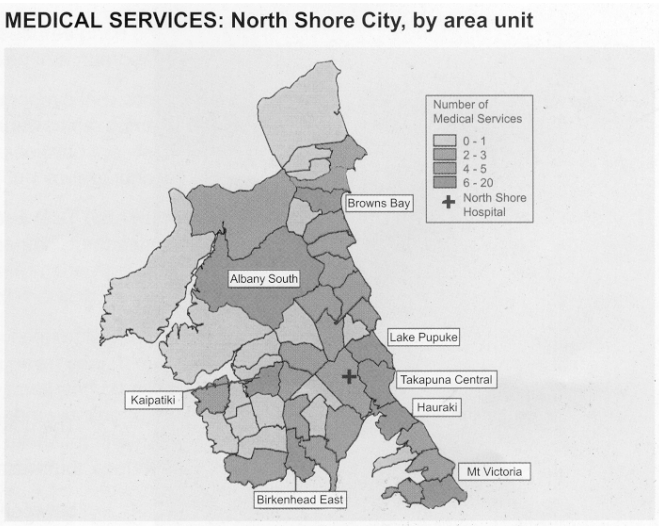
Territorial Authorities (TA). TAs are a collection of Area Units which generally define significant areas of the country and population, specifically 15 cities and 59 districts.
Area Units (AU). Area Units are aggregations of meshblocks. Area Units contain an average of 5000 persons but can vary in terms of size and population but are roughly equivalent to a suburb or part thereof.
Meshblock (MB). The meshblock is the smallest spatial unit used by SNZ in the collection and/or processing of data. It is a required field on SNZ's Business Frame and is the building block for aggregation into area units, and territorial authorities.

Uses of the Meshblock

- Enumerating, classifying and disseminating geographic information
    - locating resources (e.g. labour, raw materials).
    - locating markets (e.g. retail planning).
    - transport and routing applications (e.g. product delivery, journey to work studies)
    - civil defence and emergency planning.
    - internal migration studies.
- Linking statistical (geographic) units to higher order geographic classifications;
- Defining administrative, electoral and statistical boundaries

Example of Output

Below is an example of the kind of spatial statistic we can produce for a user. The map below shows the number and concentration of medical practices in each of the area units(suburbs) making up North Shore City. Sales territories can be defined based on the number of practices and their proximity to each other.



Limitations of the Meshblock

- **Confidentiality Issues -** meshblocks are usually aggregated to higher levels to preserve confidentiality
- **Maintenance Issues** – the meshblock pattern requires regular maintenance to reflect changing admin boundaries
- **Conceptual Issues** - takes no account of economic spatial activity
- **Operational Issues** - quality dependant on addresses supplied by respondents

Quality of geocoding on SNZ's Business Frame

SNZ carried out an investigation into the quality of the geographic information held on the frame. A geographic code, whether provided automatically or manually, was assigned to an address only if it could be guaranteed to be 100% correct. This code was then compared with the existing code on the Business Frame to determine the accuracy of current geocoding practices

Results

4.5% of the geographic unit population could not be confidently assigned a meshblock for the investigation.

79.4% of the population were correctly coded at Meshblock level
88.6% of the population were correctly coded at Area Unit level.
94.8% of the population were correctly coded at Territorial Authority level

* Due to our programs producing area unit (AU) and territorial authority (TA) information directly from a given meshblock code, we were unable to assign either an AU or TA if the meshblock was missing.

The future of geocoding in SNZ

The growing need for more precise geographic information has lead to a debate within SNZ as to whether the geocoding process best meets current and future user requirements. The question for the future of geocoding data within SNZ is: do we retain the meshblock as our lowest level spatial unit?

Issues with the current system

Conflicting Requirements - a point based system can deliver more specific information to a broader range of users ;
Quality Issues - data rarely provided at meshblock level. One reason has been concerns about the quality of the information ;
Competitive Advantage - need to be able to take full advantage of existing technologies (like GPS) to compete effectively in the marketplace ;
Comparability - SNZ's ability to match business information with census information through the existing meshblock system maximises the value of our information services ;
Confidentiality - the meshblock acts as a standard unit to protect confidentiality. Shifting to a point based system would need to be evaluated in both monetary terms and the duty of care we owe to our respondents**.**

SNZ has a statutory requirement to protect privacy and relies on the goodwill of the community for high response rates. Shifting to a point based system would need to be evaluated in both monetary terms and the duty of care we owe to our respondents. Until these issues can be resolved, the use of the meshblock coding pattern will continue to form the basis for representing spatial demographic characteristics of the NZ people and their business.